POINT SAMPLING OF LAND USE IN THE WASHITA BASIN, OKLAHOMA

April 1969 ARS 41-149
Agricultural Research Service
UNITED STATES DEPARTMENT OF AGRICULTURE

POINT SAMPLING OF LAND USE IN THE WASHITA BASIN, OKLAHOMA[1]

Windell R. Shockey and Donn G. DeCoursey [2]

Rates of soil erosion and water runoff are greatly affected by the land cover. Detailed information on plant density and height, root density and depth, extent of plant cover, and extent and amount of litter is seldom available. Therefore, it is necessary to rely largely on land use as an index of cover conditions in the hydrologic analysis of watersheds.

In 1962 the land use of 14 watersheds on 1,130 square miles of the Washita Basin in Oklahoma was mapped in detail. An observer flew over the watershed area and noted boundaries of current land use on aerial photographs. These land use areas were subsequently planimetered and divided into six categories: Alfalfa, sowed crops, row crops, open land, pasture, and timber. The cost of this inventory was $1,910-- $520 flying time and $1,390 compilation time.

A land use survey is needed annually to detect and record changes in land use. A detailed survey similar to the one made in 1962 is costly and time consuming, and may not be necessary as sampling can provide satisfactory information.

A sample can be composed of areas, lines, or points. A disadvantage of an areal sample element is the difficulty of establishing a mean value for the characteristic being evaluated within the area. Also, results tend to vary with size of the area. Line elements are sometimes difficult to obtain because of inaccessibility of points along the line. The main disadvantage of the line sample, however, is the difficulty of coding the information being obtained so that it can be analyzed. Point sampling is free of the disadvantages of the area and line samples. Furthermore, it is the easiest to obtain and analyze and, therefore, was selected as the method to be used for the land use inventory.

In point sampling, the presence or absence of the particular characteristic being inventoried is determined at preselected grid points. The prevalence of this characteristic within the area being sampled is estimated by the ratio of the number of points with the characteristic to the total number of points within the area being sampled.

A great variety of point-sampling methods have been proposed because (a) sampling methods for both of two coordinate axes may be either random or systematic; (b) the sample may be independent or stratified, and a stratified, systematic sample may be either aligned or unaligned; and (c) elements of the sample may or may not be clustered or they may be hierarchial (multistage). Discussions of the many methods of sampling can be found in Cochran, Yates, and Quenouille.[3] A general discussion of land use sampling is found in Agriculture Handbook 237.[4]

---

[1] Contribution from the Soil and Water Conservation Research Division, Agricultural Research Service, U.S. Department of Agriculture in cooperation with the Oklahoma Agricultural Experiment Station, Stillwater, Okla.

[2] Agricultural research technician and hydraulic engineer, respectively, Southern Plains Branch, ARS, USDA, Chickasha, Okla.

[3] Cochran, W. S. Sampling techniques. Wiley, N.Y. 1953; Quenouille, M. H. Problems in plane sampling. Amer. Math. Statis. 20: 355-375. 1949; Yates, F. Sampling methods for censuses and surveys. Griffin, London. 1953.

[4] Berry, B. J. L. Sampling, coding, and storing flood plain data. U.S. Dept. Agr., Agr. Handb. 237, 1962.

Both theoretical and empirical approaches show that the most efficient areal sample is the stratified systematic unaligned sample, provided the autocorrelation function of successive strata is concaved upwards. Osborne showed empirical support for upward concavity of autocorrelation functions in land use surveys.[5]

The stratified systematic unaligned sample combines the advantages of randomization and stratification with useful aspects of systematic samples, and avoids the possibility of bias caused by periodicities. It is constructed as follows (fig. 1): A square grid, of a size such that a desired number of points will fall on the area to be sampled, is placed over a map of the area. Point a is selected at random in the first stratum. The x coordinate of a is used with a new random y coordinate to locate b, a second random y coordinate to locate e, and so on across the top row of strata. By a similar process the y coordinate of a is used in combination with a random x coordinate to locate point c and all successive points in the first column of strata. The random x coordinate of c and y coordinate of b are then used to locate d. Point g is located by using the random x coordinate of f and random y coordinate of e. The same procedure is used to locate sample points in all strata.

A disadvantage of the point sample is that no dependable method for estimating the variance of the means from systematic samples is known. However, Quenouille (see reference in footnote 3) proposed three different methods for calculating the variance:

(a) Use sets of systematic samples randomly placed with respect to each other, with the error variance calculated from the variance of the systematic samples in each block.

(b) Use one set of systematic samples randomly placed, with the area then broken into blocks and the error variance calculated from the variances of the portions of the systematic samples in each block.

(c) Use one systematic sample, breaking it up into several systematic samples of wider spacing, and calculating the error variance from the portions of the subsystematic samples that fall into blocks into which the area is divided.

The three methods are increasingly accurate in their estimate of the mean, but increasingly biased in their estimate of the sample variance. However, in studies of change with time, Goodall has shown that bias can be removed from estimates of change by using the same sample for all surveys if the bias is constant through time.[6]

Cochran (see reference in footnote 3) has shown the variance of a proportion of a random sample to be

$$\sigma_s^2 = \frac{P(1.0 - P)}{N_r} \tag{1}$$

where P is the proportion of the population with a given property (a particular land use in this case) and N_r is the sample size. Thus, if one of the three methods of estimating the variance described by Quenouille were used to calculate the variance of a sample, and the value substituted into equation (1), the size of a random sample with a variance equal to the systematic sample could be found. If the size of the random sample, N_r, were compared with the sample size, N_s, the relative advantage of the systematic sample over the random sample $\frac{N_r}{N_s}$, could be calculated.

---

[5] Osborne, J. G. Sampling errors of systematic and random surveys of cover-type areas. Amer. Statis. Assoc. Jour. 37: 256-264. 1942.

[6] Goodall, D. W. Some considerations in the use of point quadrats for the analysis of vegetation. Austral. Sci. Jour., Res. Serv. B(5): 1-61. 1952.

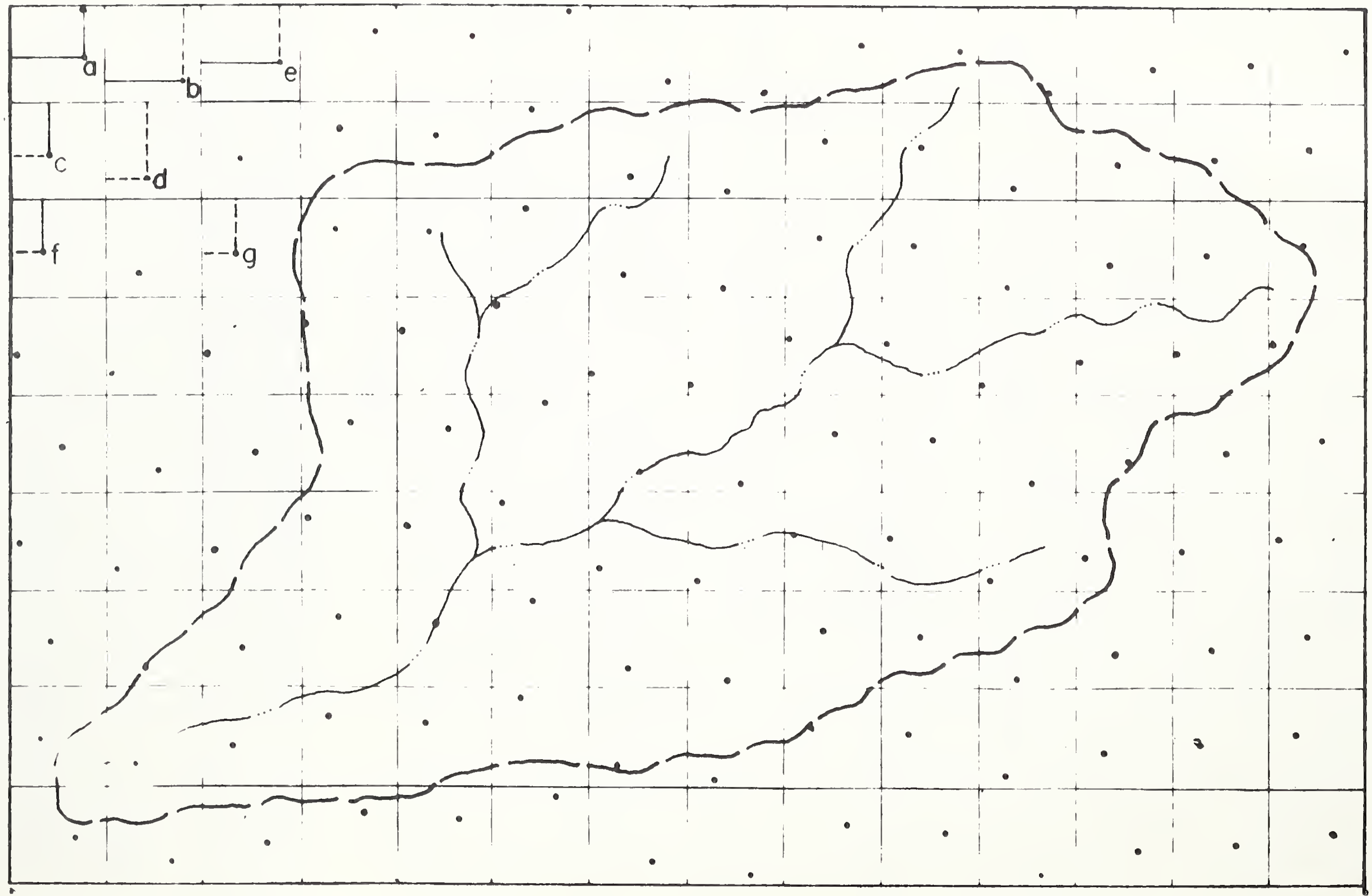

Figure 1.--Stratified systematic unaligned sample

Two 36-square-mile areas were selected for sampling and the second method of estimating the variance was used because it was easy to apply and represented a compromise between accuracy and bias. One of the areas selected for sampling was in upland or pastureland; the other was highly diversified bottom land. A square stratified systematic unaligned sample grid of 1,600 points covering 36 square miles was superimposed over aerial photographs of the area on which the land use had been marked. Land use at each grid point was noted and tabulated by a one-mile section. The grid was rotated approximately 90 degrees to give a new sample, and the land use again tabulated. The process was repeated until four sets of data had been collected on both of the areas. The land use for each set was summed at the 1-, 4-, 9-, 16-, 25-, and 36-square-mile areas, giving thirty-six 1-square-mile samples, nine 4-square-mile samples, four 9-square-mile samples, and one each of the 16-, 25-, and 36-square-mile samples.

The variance and relative advantage of the systematic over the random sample was calculated for each land use or combination thereof. Analysis of the results showed no correlation between $\frac{N_r}{N_s}$ and either the size of the subsample or the proportion of the area in a given land use, P.

Approximately 650 observations were analyzed. Figure 2 shows a plot of the number of observations per 0.25 interval or $\frac{N_r}{N_s}$. The highest concentration of points (mode) falls in the interval from 1.25 to 1.50. The midpoint of the observations (median) is 3.35. The mean of all observations $\frac{N_r}{N_s} < 30.0$ is 4.84. A value of 3.5 was selected as the relative advantage of the stratified systematic unaligned sample over the random sample. Thus the variance of the systematic sample is:

$$\sigma_s^2 = 0.286 \frac{P(1.0 - P)}{N_s}, \tag{2}$$

where σ_s^2 is the sample variance, P is the percentage of area in a given land use, and N_s is the number of systematic sample points.

Solving equation (2) for the sample size:

$$N_s = 0.286 \frac{P(1.0 - P)}{\sigma_s^2} \tag{3}$$

shows the sample size to be a function of the variance for a fixed value of P. If an allowable error, AE, is established such that it represents the 95-percent confidence limits on the sample, then the standard deviation of the sample is:

$$\sigma_s = \frac{AE}{1.96}, \tag{4}$$

where AE is the allowable error and 1.96 is the number of standard deviations within the 95-percent confidence limits. Substitution of equation (4) for the standard deviation into equation (3) gives the size of a systematic sample in terms of an allowable error:

$$N_s = 0.286 \frac{P(1.0 - P)\,(1.96)^2}{(AE)^2}. \tag{5}$$

Figure 3, based on equation (5), shows the 95-percent confidence limits of a proportion, P, of the area in a given land use for different sample sizes, N_s. For example, if a stratified systematic unaligned sample of size 200 is taken, and about 20 percent of the observations have some particular land use characteristic, the 95-percent limits are plus or minus 3 percent, and

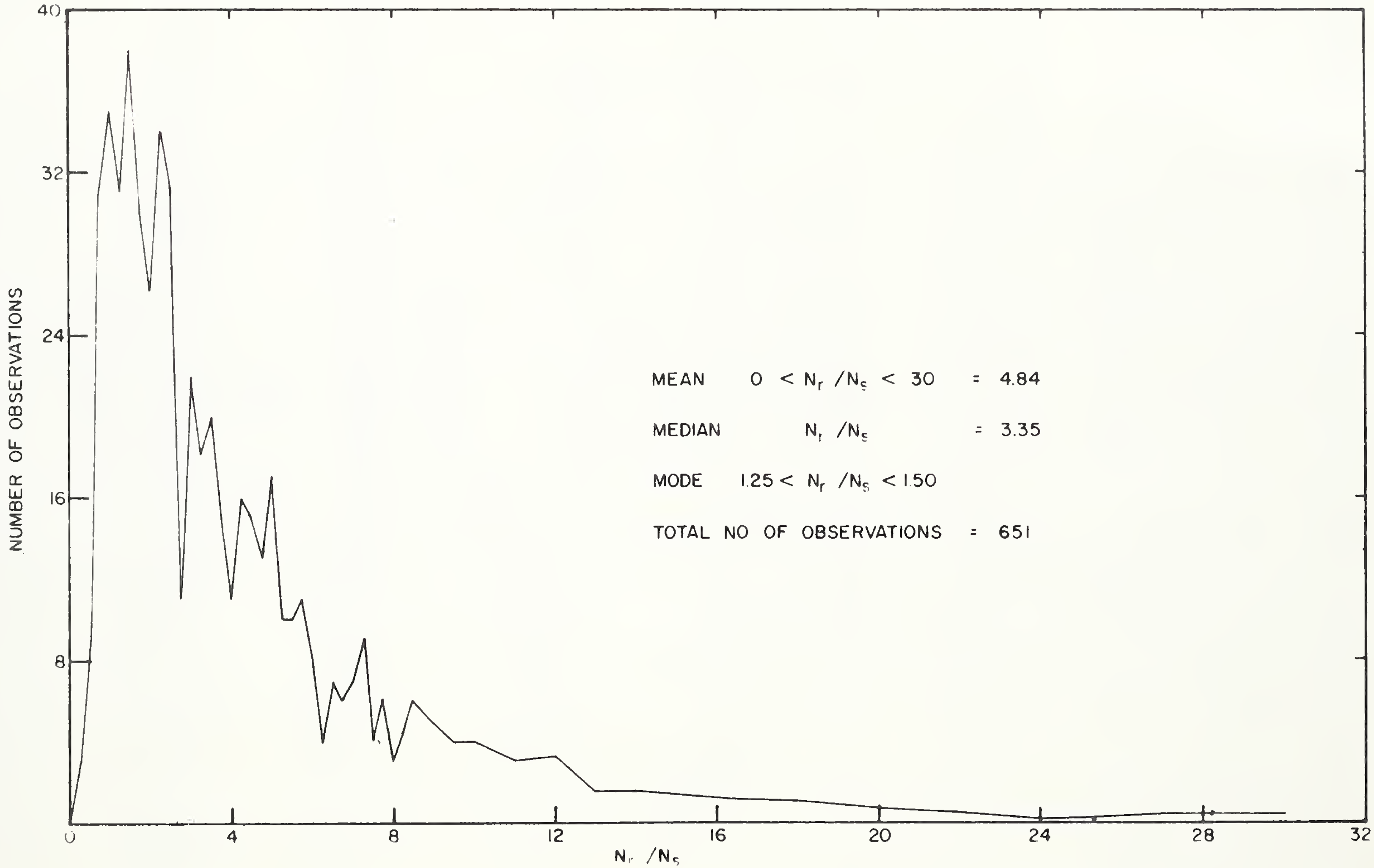

Figure 2.--Number of Observations per 0.25 interval in N_r/N_s

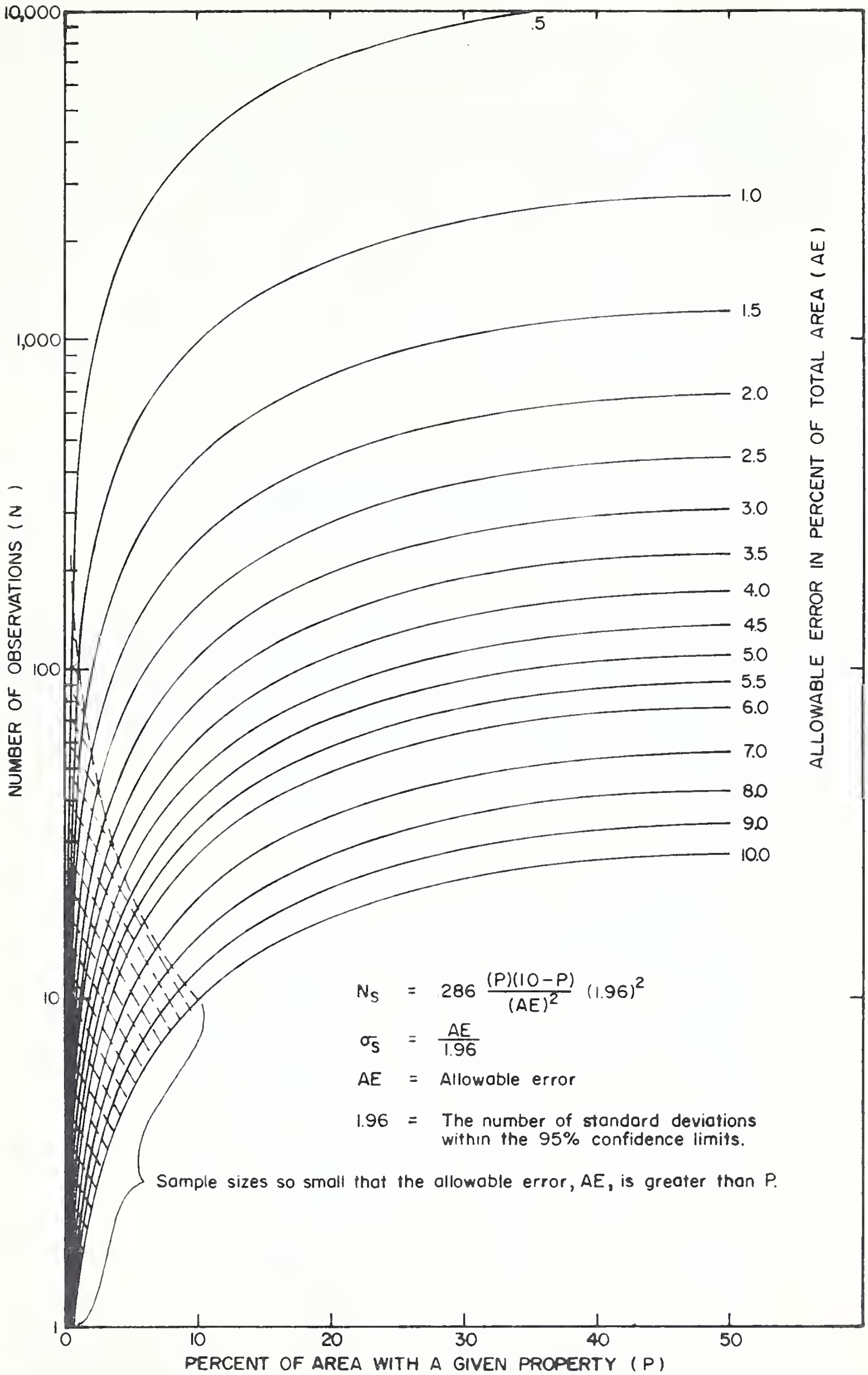

Figure 3.--95% limits of P for a stratified systematic unaligned sample

the true value of P is expected to lie between 17 and 23 percent. There is a 2.5-percent chance that P will be less than 17 and a 2.5-percent chance that it will be greater than 23-percent.

In sampling the Washita Basin, an accuracy of not less than $\pm$ 2 percent for a land use encompassing 50 percent of the area was selected as the criterion for determining the number of points to be used in the sample. From figure 3, it was found that a total of 700 points per watershed or subwatershed was required to meet this criterion. Aerial photographs were selected as base maps upon which to locate permanent observation points. The observer can locate the point and record the land use on the aerial photographs easily as he flies over the watershed. The points for sampling land use in the Washita Basin were located on the aerial photos by using an overlay. The overlay, covering 4 square miles and producing about 700 points per watershed, was constructed on graph paper. The points from the graph paper were transposed to cellulose acetate matte film paper and small holes punched at each spot. The 10-inch aerial photographs, covering 4 square miles each, were then overlayed by the grid and permanently marked in ink through the holes in the overlay.

The density of points on subwatersheds in the Washita Basin, approximately 700 points per subwatershed, varied from 36.0 points per square mile on Salt Creek to 3.65 points per square mile on Sugar Creek. A density of 25 points per square mile could be surveyed without difficulty in flying strips 1/2 mile wide. Strips 1 mile could be used on watersheds with less than 16 points per square mile.

In 1967 a complete land use inventory was made using the point sampling procedure. Figure 4 shows the study area in the Washita Basin and subwatersheds in which samples were collected. The total number of points per sample is shown on the map beside the drainage areas. The data on each watershed were tabulated by land use for each aerial photo on the form used for table 1. Thus, changes with time can be located to the nearest 4-square-mile block. The land use specifically designated in the 1967 inventory was as follows:

Alfalfa	Open ground
Row crop	Rock
Sowed crop	River or creek
Summer sowed crop	Detention dam (surface area)
Stomp lot	Farm pond (surface area)
Pasture	Highways
Gullied pasture	Farm road
Timbered pasture	Private road
Timber	Farmstead
Cleared timber	City
Gullied timber	

Table 1 is the 1967 land use inventory.

The 1967 inventory using the point sample method cost $470 -- $380 flying time and $90 compilation time. Thus, this inventory was considerably more efficient than the one in 1962.

NO	NAME	AREA (SQ MI)	NO POINTS
110	TONKAWA CREEK	39 3	WEIGHTED
	UPPER TONKAWA CREEK	26 [illegible]	[illegible]61
121	SUGAR CREEK	201 5	740
131	DELAWARE CREEK	40 [illegible]	688
141	SPRING CREEK	73 5	738
200A	WASHITA RIVER ABOVE VERDEN	72.6	719
311	SALT CREEK	23.8	880
411	LINE CREEK	53 4	696
500 A	WASHITA RIVER ABOVE TURNPIKE	167 8	726
511	WEST BITTER CREEK	60.8	703
512	EAST BITTER CREEK	35 6	685
522	LITTLE WASHITA RIVER	207 7	770
600 A	WASHITA RIVER ABOVE TABLER	75.5	688
611	BIG DRY CREEK	7 6	MAPPED
612	LITTLE DRY CREEK	.9	MAPPED
621	WINTER CREEK	33 3	814
700 A	WASHITA RIVER ABOVE ALEX	38.2	680

Figure 4.--Area of Concentration, Washita River Basin

Table 1.--1967 Land use survey in percent

Watershed Number: All Name: Watersheds Date of Survey: June 1967

Watershed No.	S	SS	A	RC	O	TG	TP	T	TC	P	GP	FS	SL	FP	DD	C	FR	PR	HW	City	Rock	Total points
110	15.9	2.1	4.6	4.7		0.5	3.7	3.0	0.6	54.0	0.4	1.2	0.1	0.5		0.4	0.1	0.3	0.6	1.6	5.7	*
111	10.9	2.6	2.0	4.1		0.8	4.5	3.1	0.6	66.8	0.6	0.5	0.1	0.5		0.5	0.1	0.4	0.1		1.8	761
121	6.5	2.7	3.1	12.7	0.1	0.1	14.3	5.0	2.6	35.7	3.1	0.7	0.1	0.5	1.5	0.7	1.4	0.1		0.3	8.8	740
131	5.4	2.2	1.7	3.9	0.6		11.9	7.0	2.3	59.3		1.2		0.1		0.7	1.2	0.7	0.2		1.6	688
141	4.7	2.6	4.2	6.4	0.1		6.1	15.3	3.7	49.2	1.1	0.4	0.4	0.4	1.0	0.5	0.3	0.3			3.3	738
200A	21.0	3.2	6.8	5.4			4.9	5.8	1.9	42.8	1.3	0.7	0.1	1.0	0.1	0.7		0.1	0.7	0.4	4.0	719
311	28.1	3.2	2.8	3.3						49.2	1.7	0.6	0.1	1.4		5.0	2.6	0.1	1.7	0.2		880
411	24.0	4.9	6.9	3.7	0.1		3.2	2.7	0.9	47.1	0.7	1.0	0.3	0.3		2.6	0.4	0.4	0.1	0.3	0.4	696
500A	28.2	5.2	7.2	3.4	0.3	0.1	1.0	0.7		41.5	2.7	0.7	0.5	1.4		1.8	1.4	0.3	0.7	2.2	0.7	726
511	17.1	4.0	3.1	4.0	1.1	1.4	1.0	0.1		57.8	4.4	1.4	0.3	0.8		0.3	1.0	0.7	1.4	0.1		703
512	4.1	1.0	3.8	0.4	0.2	2.5	1.8	0.7		75.9	6.3	0.2	0.3	1.3		0.4	0.4	0.4	0.3			685
522	9.5	2.3	2.9	2.2	0.4	0.5	7.8	5.3	1.6	52.3	5.7	1.4	0.5	0.7	0.1	1.7	1.3	0.3	1.3	0.1	2.1	771
600A	16.9	3.1	11.8	5.5	0.2	0.6	2.6	1.6	0.6	45.3	4.5	0.7	1.2	0.3		2.0	1.0	0.4	1.0	0.7		688
621	3.7	1.0	2.5	1.1	0.3	1.5	7.3	1.5	1.5	66.6	7.1	0.9	0.4	1.3	0.5	1.0	0.9	0.4	0.5			814
700A	8.5	2.8	6.9	9.3		0.2	5.0	1.8		56.1	4.6	0.4	0.2	0.4	0.3	1.8	0.1	0.9	0.3	0.4		680

*Weighted.

Key:
- A-Alfalfa
- C-Creeks
- DD-Detention dams
- FP-Farm pond
- FR-Farm road
- FS-Farmstead
- GP-Gully pasture
- HW-Highways
- O-Open ground
- P-Pasture
- PR-Private road
- RC-Row crop
- S-Sowed crop
- SL-Stomp lot
- SS-Summer sowed
- T-Timber
- TC-Timber cleared
- TG-Timber gully
- TP-Timber pasture

The point-sample method was applied to the 1962 land use data to compare the point-sample results with the detailed survey results. Table 2 is a summary of the land use by both methods for each subwatershed. The results are also plotted in figure 5. If it is assumed that there is no error in the planimetered data, the inner lines represent the 95-percent confidence limits used in setting up the survey. If a 2-percent relative error is allowed in the planimetered data, the outer lines are the 95-percent confidence limits on the system. Figure 5 indicates that the point-sampling method is at least as accurate as the planimetering method and possibly more accurate if the criterion used in establishing the number of points is correct.

Table 2.--Comparison of planimetered and point samples, 1962 land use survey, Washita River Basin

Number and name of watershed	Sowed crop		Row crop		Alfalfa		Open		All other	
	Plan.	Point	Plan.	Point	Plan.	Point	Plan.	Point	Plan.	Point
	Pct.	Pct.	Pct.	Pct.	Pct.	Pct.	Pct.	Pct.	Pct.	Pct.
110 Tonkawa Creek	14.2	15.4	7.2	7.4	4.1	4.1	0.2	0.5	74.3	72.6
111 Upper Tonkawa Creek	10.0	10.4	7.6	7.9	2.9	2.9	0	0	79.6	78.8
121 Sugar Creek	6.7	6.2	17.7	15.0	1.6	1.5	T	0	74.0	77.3
131 Delaware Creek	6.8	6.4	9.5	9.3	.7	.4	T	0	83.0	83.9
141 Spring Creek	4.1	4.2	9.6	8.6	1.9	2.0	.1	.1	84.3	85.1
200A Washita River above Verden	21.1	21.7	11.1	11.4	6.2	7.2	.8	1.3	60.8	58.4
311 Salt Creek	23.3	23.8	4.2	3.7	4.3	3.9	1.3	1.3	66.9	67.3
411 Line Creek	21.9	23.9	6.2	6.6	8.6	7.5	.2	.1	63.1	61.9
500A Washita River above Turnpike	27.3	25.3	6.9	4.4	8.0	7.6	.5	.3	57.3	62.4
511 West Bitter Creek	13.7	13.6	6.8	7.3	3.2	3.4	.4	.1	75.9	75.6
512 East Bitter Creek	5.3	4.1	1.7	1.5	2.6	2.3	.1	0	90.3	92.1
522 Little Washita River	10.8	11.5	3.4	3.8	2.6	2.0	.3	.5	82.9	82.2
600A Washita River above Tabler	17.2	19.7	7.1	6.7	9.8	8.6	.9	.7	65.0	64.3
621 Winter Creek	3.8	3.7	1.0	1.3	2.8	3.6	.6	.2	91.8	91.2
700A Washita River above Alex	7.3	6.9	10.7	10.0	8.9	9.3	.4	.3	72.7	73.5

CONCLUSION

In 1962 a land use inventory of 1,130 square miles of the Washita Basin cost $1,910. Since land use surveys may be made annually, a more economical method was needed. Point sampling was selected as the best method, with a stratified systematic unaligned type of sample as the most efficient. Tests of this type of sample showed it was about 3.5 times as efficient as random samples. A land use inventory made in 1967 based on a systematic sample cost $470, and results were at least as accurate as the 1962 survey.

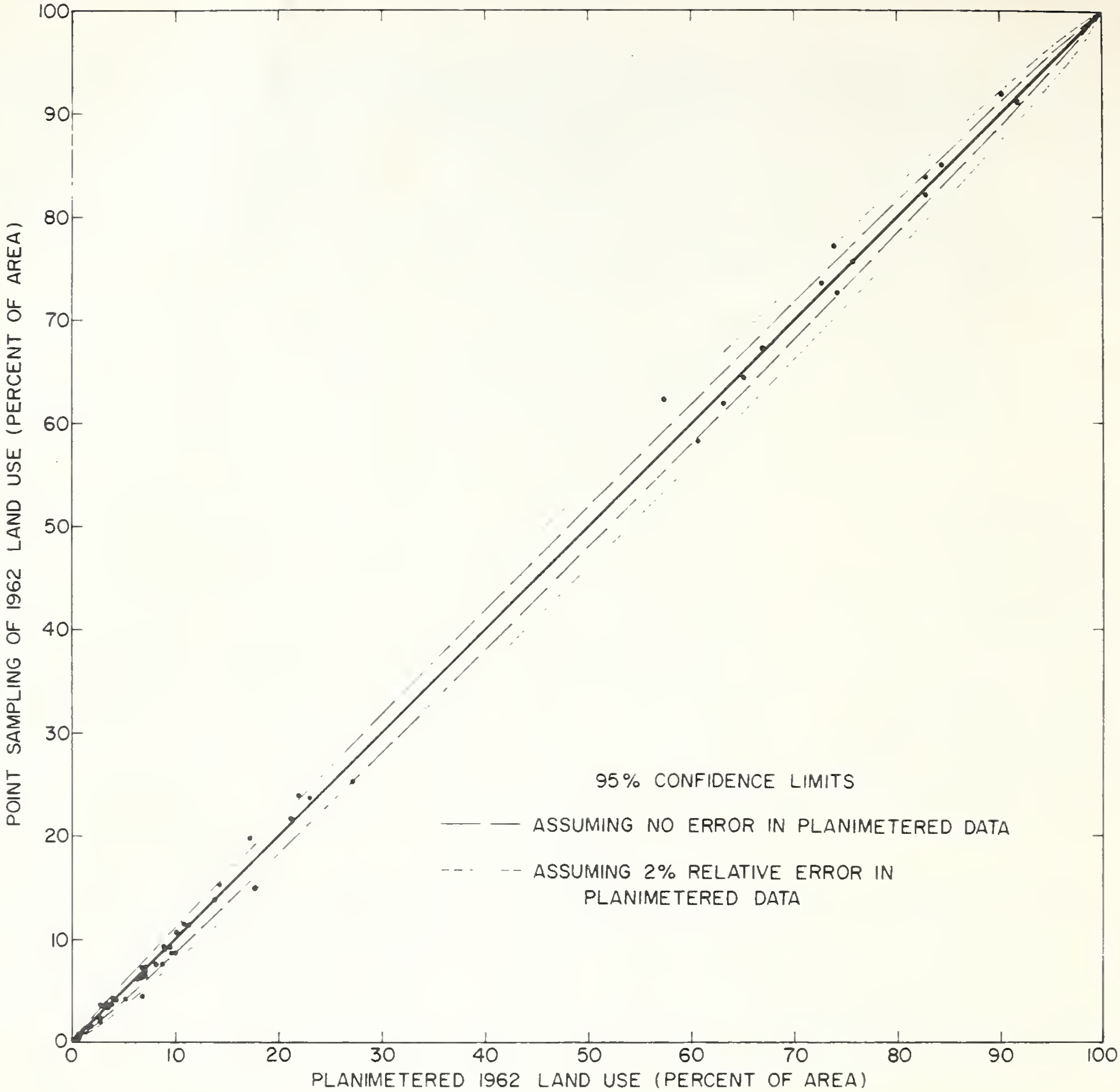

FIGURE 5.--SAMPLED VS. PLANIMETERED, 1962 LAND USE

Figure 5.--Sampled vs. Planimetered, 1962 Land Use